

INTELLIGENT DATA MANAGEMENT FOR REAL-TIME SPACECRAFT MONITORING

Ursula M. Schwuttke

Jet Propulsion Laboratory
California Institute of Technology
4800 Oak Grove Dr., Pasadena, CA 91109
ursula@nereid.jpl.nasa.gov

Les Gasser and Bruce Abramson

Computer Science Department
University of Southern California
Los Angeles, CA 90089
gasser@usc.edu

Abstract

Real-time AI systems have begun to address the challenge of restructuring problem solving to meet real-time constraints by making key trade-offs that pursue less than optimal strategies with minimal impact on system goals. Several approaches for adapting to dynamic changes in system operating conditions are known. However, simultaneously adapting system decision criteria in a principled way has been difficult. Towards this end, a general technique for dynamically making such trade-offs using a combination of decision theory and domain knowledge has been developed. The paper discusses multi-attribute utility theory (MAUT), a decision theoretic approach for making one-time decisions, describes dynamic trade-off evaluation as a knowledge-based extension of MAUT that is suitable for highly dynamic real-time environments, and provides an example of dynamic trade-off evaluation applied to a specific data management trade-off in a real-world spacecraft monitoring application.

1. Introduction

Lengthy response times often prohibit optimal problem solving in the presence of real-time constraints. Effective real-time systems therefore require meta-reasoning for making appropriate trade-offs and, when necessary, pursuing less optimal methods. Such meta-reasoning often must take place in the presence of incomplete information, insufficient resources, and unpredictable situations; precise mathematical approaches with parallels in traditional control theory therefore cannot be formulated. As a result, the applicability of decision theory and the psychology of judgment to this problem area was recognized early, with research on heuristic methods for inference control [Simon 1955]. However, initial enthusiasm for using decision theory as an artificial intelligence technique dwindled in favor of approaches that seemed to lend themselves more naturally to expressing the rich structure of human knowledge [Horvitz 1988].

Uncertainty in reasoning has since been expressed with probabilities and statistics and has been thoroughly explored

for nonreal-time AI applications in the context of Bayesian belief representation [Pearl 1988], [Shachter 1987]. However, degrees of uncertainty in real-time situations can change rapidly, imposing overwhelming complexities on these techniques. Bayesian statistics relies on the availability of conditional probabilities for the various hypotheses and pieces of evidence that pertain to a given situation. A common implementation of Bayesian statistics appears in medical expert systems that calculate the probability that a patient is suffering from a particular disease, given a manifested set of symptoms. Even in this relatively simple application with slow trend changes, the prospect of deriving the needed statistics is not straightforward: it is beset by a multitude of questions and variables such as when to use global statistics rather than local ones, how often to update models to reflect changing trends, and, more fundamentally, how to get access to valid information given inaccuracies and varying procedures in disease reporting. A central difficulty associated with the use of Bayesian statistics is therefore centered around dependence on stable information about the domain environment: this information, which is difficult to obtain even in simpler real-world situations, can be impossible to derive dynamically for complex real-time problems.

For such reasons, there has been renewed interest in decision theory for real-time AI applications. Rapidly changing circumstances require making trade-offs and expressing judgments, two processes which can entail a substantial level of subjectivity [von Winterfeldt 1986] and are therefore incompatible with rigid methods of analysis that require stable and accurate information. Decision theory provides a key ingredient: flexibility. This flexibility is embodied in formal decision-theoretic principles for obtaining preferred courses of action in the presence of uncertain events and conflicting objectives.

The simultaneous consideration of time pressure, complex environments, and potentially conflicting objectives has been studied in several different settings, including game playing [Russell 1989] and medical decision-making [Hor-

vitz 1989]. Studies of rational agents with time constraints have also begun to emerge in the literature [D'Ambrosio 1990], [Hansson 1990]. Most of these studies, however, rely upon assumptions that are not universally applicable. Both game-playing and medical diagnosis, for example, are self-contained domains with relatively long time lapses allowed between stimulus and response. Many interesting domains, such as spacecraft monitoring, violate both assumptions: the number of potential variables is huge and response time is negligible. The problem of automated spacecraft monitoring provides a perfect example of a real-time environment that is characterized by changing circumstances, uncertain events, and conflicting computational objectives. Several real-time systems with knowledge-based components have been developed for this complex application domain, [Laffey 1988], [Muratore 1990], [Schwutke 1990] but these systems have focused primarily on being fast enough to handle expected computational loads and not on responding dynamically to unforeseen changes in real-time demands.

2. Dynamic Trade-Off Evaluation

Multi-attribute utility theory offers a natural way for dealing with competing objectives and is computationally straightforward, but has not been applied in dynamic real-time environments. Although a variety of static techniques from multi-attribute utility theory exist, only three variants of these techniques have been commonly applied to real-world situations [von Winterfeldt 1986]: the simple, multi-attribute rating technique [Edwards 1977], difference value measurement [Dyer 1979], and subjectively expected utility (SEU) measurement [Keeney 1976]. These approaches consist of the same general procedures and have collectively become known as SEU techniques. Edwards' technique is not only the simplest computationally, but also the most amenable to combination with knowledge-based approaches. It has thus been selected for our extension to dynamic real-time environments.

Our approach to this extension is to modify the basic SEU procedures while attempting to maintain their inherent simplicity, robustness, and flexibility. Our procedure is termed Dynamic Tradeoff Evaluation (DTE). In DTE, utility theory is used to rank alternatives in a preference space, and knowledge-based decision rules are used at run-time to 1) dynamically re-weight the attributes of individual alternatives and 2) to dynamically select among preference criteria in the preference space (depending on situational attributes and operational mode). The DTE methods are sufficiently general that they are applicable to a variety of run-time trade-offs, and are currently being applied to several very different real-time, real-world trade-offs in the domain of spacecraft monitoring. (See [Schwutke 1991], which discusses and classifies a large range of potential trade-offs in Real-time AI.) The DTE methods are sufficiently general that they are applicable to a variety of run-time trade-offs and to integration into a real-time monitoring architecture.

The DTE procedure involves a sequence of six steps, many of which are derived from the steps of static SEU procedure. The first three of these steps and part of the fourth must be completed during the design phase of the system. For a given trade-off, the procedure includes:

1. *Definition of the trade-off instantiation mechanism.* This step involves specifying the circumstances under which DTE is required and designing the mechanism that will detect those circumstances and invoke the trade-off evaluation.

2. *Definition of application-specific alternatives and criteria that determine the value of the alternatives.* During this step, the alternative actions to be considered in the trade-off evaluation are specified, along with criteria that will be used to evaluate the alternatives. As part of this process, the system designers and domain experts also specify domain knowledge and (if necessary) heuristics that define the various ways of implementing each alternative. In addition, the decision criteria that influence the specific implementation of a run-time alternative are considered.

3. *Separate evaluation of each alternative.* This is done in conjunction with the previous step, and involves reliance on subjective judgements in cases where no basis for objective evaluation exists. Each alternative is ranked with respect to each of the evaluation criteria, on a scale of 0 to 100, and suitable consistency checks are applied to the evaluation.

4. *Definition of weights and modes.* Relative weights are assigned to each of the criteria, along with ranges within which the weights can vary. Domain knowledge is specified to determine the circumstances under which the weights will be varied. In addition, multiple modes may be specified, where each mode is governed by a different set of weights. Both the variation of the weights and the choice of a mode are determined at run-time using domain knowledge. These decisions are based on instantaneous circumstances in the monitored environment.

5. *Aggregation.* The weights selected in the previous step are used to determine the aggregate value of each of the alternatives, using the additive aggregation model put forth in SEU. These aggregate values provide the evaluation of the alternatives with regard to one of the trade-off axes. Depending on the specific trade-off, similar evaluation and aggregation may be required with regard to the second trade-off axis. However, in many cases the evaluation on the second axis may be directly calculated based on the dynamics of the environment. (In applications that do not require domain knowledge, the evaluations on both axes may be directly obtainable, but these applications are considered peripheral to this research.)

6. *Selection.* An alternative is selected based on greatest total value with respect to both trade-off axes, as specified in the SEU methods. When the evaluation indicates that two or more alternatives are equally good, domain knowledge is used to select one alternative over the others, or if the alternatives are not mutually exclusive, to select several of them.

3. Application: Telemetry Data Management

We describe the application of DTE by reference to its application in a real-world spacecraft monitoring problem: managing input data for real-time knowledge-based monitoring of telemetry data from the Galileo Solid-State Imaging (SSI) system.

The basic real-time task for mission operations involves comparing incoming engineering telemetry to a combination of predicted data values and limit ranges. Specific predictions reflect subsystem goals that result from the planned sequence of subsystem events, and the limit ranges reflect the general operating parameters of the instrument. This task involves two AI components: intelligent input data management and knowledge-based anomaly detection/analysis, in addition to the basic real-time monitoring task. Here we focus on the first of these. The (competing) goals of intelligent data management in this application are to dynamically adjust input data volumes to meet the processing capabilities of the host hardware, while maximizing the information content, maintaining alertness to unusual events in the input data, and focusing on particularly relevant tasks. The particular trade-off we examine in this paper to illustrate our technique is a timeliness trade-off: representativeness of the input data versus timeliness of the solution.

In SSI, four possible data management alternatives have been specified as a result of extensive interviews with an imaging subsystem specialist as part of the first step of DTE. These alternatives are: eliminating channels not in the basic monitoring set, eliminating channels not in the minimal set, reducing sampling rate on heuristically defined subset of channels, and reducing sampling rate on the entire channel set. The converse set of alternatives applies when data rates or computational load from other processes decrease. These converse alternatives include adding channels in the full monitoring set, adding channels in the basic set, increasing sampling rate on a selected channel subset, and increasing sampling rate on the entire channel set.

The four specified alternatives (numbered 1.1, 1.2, 2.1, and 2.2 respectively) are evaluated with regard to three criteria that define data representativeness. For data reduction, these include: (A) non-dynamic behavior, (B) irrelevance to an existing problem area, and (C) non-negative impact on monitoring integrity. A data channel must exhibit non-dynamic behavior before it can be eliminated; frequent channel value changes indicate a high level of activity that must be monitored to maintain adequate representativeness. When representativeness is an issue, irrelevance to existing problem areas is important in deciding which channels to remove from the monitored set. Finally, only channels that do not compromise monitoring process integrity in current circumstances can be eliminated without impacting representativeness. Conversely, when the size of the monitoring set is being increased, the criteria must become (A)

dynamic behavior, (B) relevance to an existing problem area, and (C) positive impact on monitoring integrity.

The second step also requires the specification of domain knowledge that shows how to implement the alternatives. In SSI, the channel elimination alternatives and the second sampling rate alternative are influenced most heavily by a decision tree that defines deletable channel subsets and the circumstances under which they apply. There are also exceptions that apply to some deletable subsets with respect to criterion (A). This exception arises because channels with a significant level of activity should not be eliminated from the monitored set even if they are part of an appropriate deletable subset. In contrast, the heuristically-defined sampling rate alternative is entirely governed by the specific situation in which it is applied. In a *normal operating mode*, the sampling rate can be reduced on all channels that are not part of the critical subset. In an *anomaly detection mode*, the sampling rate should only be reduced on channels that are irrelevant to anomaly detection. However, in the event of large backlogs, reduction on sampling of all channels may be desirable.

Occasionally channels must be added irrespective of timeliness. This is because in anomaly detection mode, increased representativeness takes instant precedence, and channels pertinent to that anomaly are added. With multiple simultaneous anomalies, more channels may be needed. Subsequently, timeliness considerations may be applied to some of the other channels in the monitoring set. When the system returns to a normal operating mode, the channels relevant to a previously resolved anomaly may be candidates for removal from the monitoring set if timeliness must be improved.

In the third step, relative weights are assigned to the attributes. Initial weights and variance ranges for these weights are defined so the weights can be adjusted during the reasoning process. This allows the weights to accommodate changing circumstances in the monitored environment. Weight variations are initiated when the system detects that its performance is degrading, and are implemented using rules that provide updates based on situational parameters. This step also entails subjectively ranking each alternative in the context of each criterion at design time, as shown in Figure 1. The ranking, obtained and checked for consistency with the help of the subsystem expert, is on a scale of 0 to 100 (with 100 having the maximum value). For example, alternative 2.1 obviously ranks the highest with regard to B, because the expert specifically designed this alternative not to impact channels with relevance to an existing problem area. Alternative 1.1, which removes the largest number of anomaly-related channels, is perceived to be the poorest choice with regard to criterion B. Conversely, when judged against criterion C, alternative 1.1 has the highest ranking because the channels that it removes generally are the first to be removed and are only added back in small subsets in the event of anomalies. Two

sets of weights are defined for this application, as shown in Figure 2. The first set applies in normal operating mode and the second applies in anomaly detection mode. In normal operating mode, irrelevance of a channel to an existing problem area is given no weight, because no problems are present. However, in anomaly analysis mode, this attribute receives the greatest weight.

ATTRIBUTE	ALTERNATIVE NUMBER			
	1.1	1.2	2.1	2.2
A	75	90	30	40
B	20	30	90	50
C	100	75	40	25

Figure 1. Values of the Alternatives for the Galileo SSI Trade-off.

In the fourth step, the single-attribute alternative rankings and the attribute weights are aggregated into an overall evaluation of alternatives which combines with the application-specific domain knowledge to enable the selection of most valuable alternative for the given circumstances. This step differs most significantly from the comparable static SEU step for two reasons. First, circumstances dictate varying weights, which in turn dictate varying aggregations. Secondly, circumstances may vary the knowledge that is applied from situation to situation. Examples of the varying aggregations that are obtained for both operating modes are shown in the tables of Figure 3. These tables show that the data management actions that are most compatible with maintaining maximum representativeness are determined by external circumstances. The ranking of the four alternatives with regard to representativeness value in varying circumstances is summarized in Figure 4, with 1 being the highest ranking and 4 being the lowest.

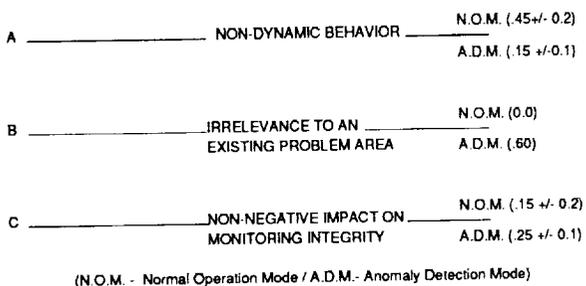


Figure 2. Attributes and Weights for Telemetry Reduction in the Galileo SSI Trade-off.

The final activity of this procedure involves the selection of an alternative based on dynamic evaluation of the representativeness vs. timeliness trade-off. In order to make this

trade-off, the four alternatives must also be evaluated with regard to timeliness. The timeliness impact of an alternative is directly proportional to the percentage reduction (or increase) in the number of monitored channels that results from implementing that alternative. However, this percentage must be calculated immediately prior to making the trade-off, based on the number of channels in the current set, because the number of monitored channels is a dynamic quantity determined by the set of events leading up to the current circumstances. To clarify the dynamic and adaptive nature of this evaluation, we consider the following example.

Assume that the monitoring system has just been brought on-line. Initially, all 49 channels are in the monitored set. After some time the system detects that an input backlog is building, and responds by deciding that some channels must be removed from the monitored set. No anomalies have been detected as yet, and no modifications to the starting weights have been suggested by the knowledge base. As a result of this situation, the system finds itself using the aggregate values in the first line of Figure 3 (top) as representativeness values.

Timeliness values are obtained by calculating the net percentage reduction in input data. Alternative 1.1 eliminates the channels not in the minimal set, or 32 of the 49 channels. Alternative 1.2 eliminates the channels not in the basic set, or 24 of the 49 channels (as governed by domain rules that are not discussed in detail here). These alternatives therefore result in a 65% and a 50% reduction respectively. According to our heuristics, the reduced sampling alternatives can eliminate 4 out of every 5 input values when no anomalies are present. Thus, with alternative 2.1, we can eliminate 80% of a subset of the monitored set. Under the present circumstances, this subset consists of all channels not in the basic set. A reduction of 80% is therefore possible on 24 of the 49 channels. With alternative 2.2, we eliminate 80% of the sampling on the entire channel set, resulting in reductions of 50% and 80% respectively. The percentage reductions are plotted against the aggregate representativeness value for each alternative as shown in Figure 5 (left). Both representativeness and timeliness are thus rated on a scale of 0-100; one unit on the representativeness scale is equivalent to one unit on the timeliness scale. The indifference curves shown in the figure are implied by this constant trade-off of units, alternatives lying on the same indifference curve have equivalent value, and alternatives lying nearest to the upper right of the graph are perceived as best. For this application, the alternatives in order of preference are 1.1, 1.2, 2.2 and 2.1. (Note that timeliness considerations have changed the order of preference from that shown in Figure 4, which is based on representativeness alone.) As a result of this analysis, alternative 1.1 is selected and implemented. Our system is now actively monitoring only 17 of the 49 channels, and is achieving adequate throughput. We will assume that at some later time, an anomaly appears on channel 1910, which requires three additional channels to be added.

Attribute	Weight*	Weight**	Weight***	ALTERNATIVE NUMBER			
				1.1	1.2	2.1	2.2
A	0.45	0.65	0.25	75	90	30	40
B	0	0	0	20	30	90	50
C	0.55	0.35	0.75	100	75	40	25
Aggregate Value* (using weight*)				88.57	81.75	35.5	31.75
Aggregate Value** (using weight**)				83.75	84.75	33.5	34.75
Aggregate Value*** (using weight***)				93.75	78.75	37.5	28.75

- * N.O.M. with no modification on starting weights
- ** N.O.M. with weight modification for greater emphasis on environmental dynamics
- *** N.O.M. with weight modification for greater emphasis on overall monitoring integrity

Attribute	Weight*	Weight**	Weight***	ALTERNATIVE NUMBER			
				1.1	1.2	2.1	2.2
A	0.15	0.25	0.05	75	90	30	40
B	0.6	0.6	0	20	30	90	50
C	0.25	0.15	0.35	100	75	40	25
Aggregate Value* (using weight*)				48.25	49.75	68.5	42.25
Aggregate Value** (using weight**)				45.75	51.75	67.5	43.75
Aggregate Value*** (using weight***)				50.75	48.75	69.5	33.25

- * A.D.M. with no modification on starting weights
- ** A.D.M. with weight modification for greater emphasis on environmental dynamics
- *** A.D.M. with weight modification for greater emphasis on overall monitoring integrity

Figure 3. Aggregate Values of Alternatives for Varying Weights in Normal Operational Mode (top) and Anomaly Detection Mode (bottom).

MODE \ ALTERNATIVE	Elimination of chan. not in basic subset	Elimination of chan. not in critical subset	Sampling reduction on heuristic subset	Sampling reduction on entire subset
N.O.M. with no modification	1	2	3	4
N.O.M. with backlog modification	2	1	4	3
N.O.M. with monitoring modification	1	2	3	4
A.D.M. with no modification	3	2	1	4
A.D.M. with backlog modification	3	2	1	4
A.D.M. with monitoring modification	2	3	1	4

Figure 4. Rankings of data management alternative values with respect to representativeness for various modes.

The anomaly is solved, and at some later time, another anomaly appears on channel 1881, requiring the addition of 12 more channels.

We are now actively monitoring 32 channels, and are beginning to build a backlog. The system's backlog detection module detects the backlog, and initiates meta-reasoning to reduce it. Figure 5 (right) shows the re-evaluation in response to the environment change at this point. The analysis will have been as follows. Alternatives 1.1 and 1.2 will both allow only 3 channels of the 32 channels in the monitored set to be eliminated. This is because 12 of the channels per-

tain to the current anomaly and 17 belong to the minimal set. Thus, both alternatives achieve a 9.3% reduction in input data. Alternative 2.1 reduces sampling on approximately 60% of the channels in the sampling set, but because we are in the anomaly detection mode, we now only filter half of the input data from these channels, achieving an effective reduction of 30%. With alternative 2.2, we filter half of the input data on all 32 channels for an effective reduction of 50%. These values are plotted against representativeness as shown in Figure 5. In this case, however, the selection of an alternative is not as obvious as in the previous iteration; alternatives 2.1 and 2.2 are very close to lying on the same

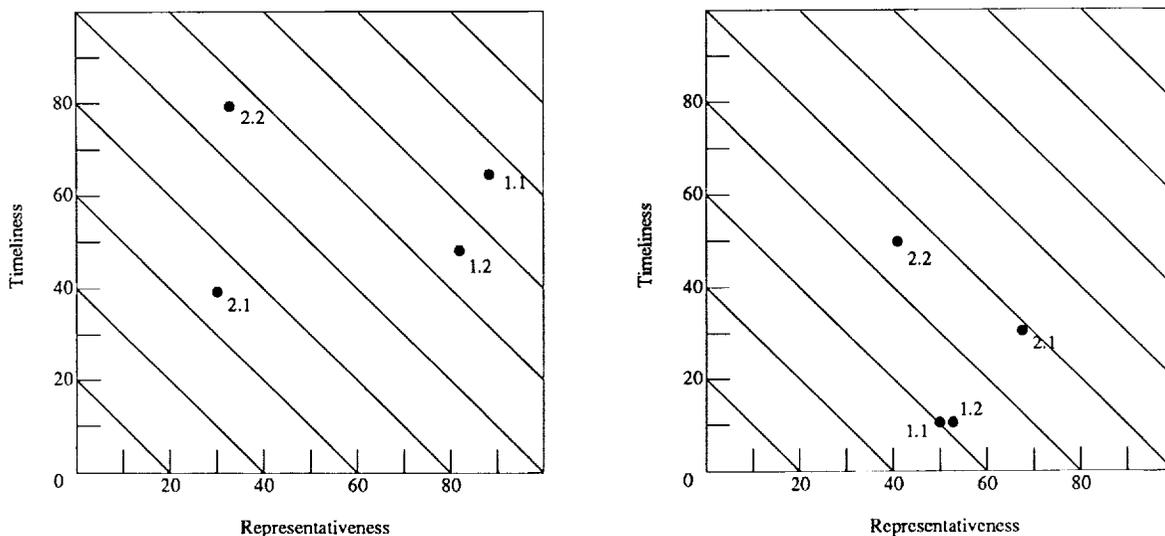


Figure 5. Timeliness vs. Representativeness for Dynamic Input Data Management in Galileo SSI Monitoring. Figure (a), at the right, shows the evaluation of the tradeoff that is made at system initialization, and Figure (b), at the left, shows the how the dynamics of the environment change after several hypothetical anomalies have been detected.

indifference curve. However, heuristics indicate that in the anomaly detection mode, representativeness is the more important consideration, and alternative 2.1 must be selected. Eventually, the anomaly on channel 1881 is resolved, and we return to the normal operation mode. Assuming no change in data rate, in this mode a similar analysis will cause the system to return to its original choice of alternative 1.1, and to continue fully monitoring only channels in the basic subset.

This example has shown the effectiveness of combining decision theory with heuristics to dynamically make real-time trade-offs for intelligent data management. The example illustrates the dynamic nature of the decision environment, and demonstrates the ability to use domain specific heuristics to guide the trade-off process and achieve real-time meta-reasoning for run-time control.

4. Conclusions

Several approaches are known for adapting AI problem solving to dynamic changes in system operating conditions, but simultaneously adapting decision criteria in a principled way has been difficult. This paper has described a general technique for dynamically making performance trade-offs to achieve these ends using a combination of decision theory and heuristic domain knowledge. Dynamic Tradeoff Evaluation is a knowledge-based extension of multi-attribute utility theory. In DTE, multi-attribute utility theory is used to rank alternatives in a preference space and heuristic decision rules are used at run-time to dynamically

re-weight the attributes that govern the value of individual alternatives. This enables dynamic selection among preference criteria in the preference space, depending on situational attributes and operational modes. DTE is suitable for highly dynamic real-time environments, as illustrated by its application in specific trade-offs spacecraft telemetry monitoring. It provides a new, rigorous, and effective way to simultaneously adapt system decision criteria and problem-solving parameters.

5. Acknowledgment

The research described in this paper was carried out in part at the Jet Propulsion Laboratory, California Institute of Technology under a contract with the National Aeronautics and Space Administration. The domain knowledge pertaining to mission operations and solid state imaging that was used in this research was contributed by William Cunningham at the Jet Propulsion Laboratory.

6. References

- D'Ambrosio 1990] B. D'Ambrosio. Constrained Rational Agency. To appear in IEEE Transactions on Systems, Man, and Cybernetics. 1990.
- [Dyer 1979] J. S. Dyer and R. A. Sarin. Measurable Multi-attribute Value Functions. Operations Research, 22, 810-822.

- [Edwards 1977] W. Edwards. How to Use Multi-attribute Utility Measurement for Social Decision Making. IEEE Transactions on Systems, Man and Cybernetics, SMC-7, 326-40.
- [Horvitz 1988] E. J. Horvitz, J. S. Breese, and M. Henrion. Decision Theory in Expert Systems and Artificial Intelligence. International Journal Of Approximate Reasoning, Special Issue on Uncertain Reasoning: 247-302, 1988.
- [Horvitz 1989] E. J. Horvitz, G. F. Cooper, and D. E. Heckerman. Reflection and Action Under Scarce Resources: Theoretical Principles and Empirical Study. Proceedings of the International Joint Conference on Artificial Intelligence. Detroit, MI 1989.
- [Keeney 1976] R. L. Keeney and A. Sicherman. An Interactive Computer Program for Assessing and Using Multi-attribute Utility Functions. Behavioral Science, 21, 173-182.
- [Laffey 1988] T. J. Laffey, S. Weitzenkamp, J. Read, S. Kao, and J. Schmidt. Intelligent Real-time Monitoring. Proceedings of National Conference on Artificial Intelligence, 1988.
- [Muratore 1990] J. F. Muratore, T. A. Heindel, T. B. Murphy, A. N. Rasmussen, R. Z. McFarland. Acquisition at Mission Control. Communications of the ACM, 33:12. December 1990.
- [Pearl 1988] J. Pearl. Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Morgan Kaufmann Publishers, San Mateo, CA., 1988.
- [Russell 1989] S. J. Russell and E. H. Wefald. On Optimal Game-Tree Search Using Rational Meta-reasoning, International Joint Conference on Artificial Intelligence, Detroit, Michigan, 1989, 334-340.
- [Shachter 1987] R. D. Shachter and D. E. Heckerman. Thinking Backward for Knowledge Acquisition. AI Magazine, 8, 55-63, 1987.
- [Schwutke 1990] U. M. Schwutke, J. R. Veregge, R. Angelino, C. L. Childs. MARVEL: An Automated Productivity Enhancement Tool for Multi-mission and Multi-subsystem Spacecraft Operations. First International Symposium on Ground Data Systems for Spacecraft Control. Darmstadt, W. Germany. 1990.
- [Schwutke 1991] U. M. Schwutke. Intelligent Real-time Monitoring of Complex Systems. Ph.D. Thesis in progress. University of Southern California., Los Angeles, CA. 1991.
- [Simon 1955] H. A. Simon. A Behavioral Model of Rational Choice. Quart. J. Econ., 69, 99-118.
- [von Winterfeldt 1986] D. von Winterfeldt and W. Edwards. Decision Analysis and Behavioral Research. Cambridge University Press. 1986.